



Published in final edited form as:

Cell. 2016 October 6; 167(2): 325–339. doi:10.1016/j.cell.2016.08.031.

## Field Guide to Plant Model Systems

Caren Chang<sup>1,\*</sup>, John L. Bowman<sup>2,\*</sup>, and Elliot M. Meyerowitz<sup>3,\*</sup>

<sup>1</sup>Department of Cell Biology and Molecular Genetics, University of Maryland, College Park, MD 20742-5815, USA

<sup>2</sup>School of Biological Sciences, Monash University, Clayton Campus, Melbourne, Victoria 3800, Australia

<sup>3</sup>Howard Hughes Medical Institute and Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA 91125, USA

### Abstract

For the past several decades, advances in plant development, physiology, cell biology, and genetics have relied heavily on the model (or reference) plant *Arabidopsis thaliana*. *Arabidopsis* resembles other plants, including crop plants, in many but by no means all respects. Study of *Arabidopsis* alone provides little information on the evolutionary history of plants, evolutionary differences between species, plants that survive in different environments, or plants that access nutrients and photosynthesize differently. Empowered by the availability of large-scale sequencing and new technologies for investigating gene function, many new plant models are being proposed and studied.

### Introduction

*Arabidopsis thaliana* has been the major plant model system in the past three decades, allowing for tremendous advances in our understanding of plant development, signaling, hormone biology, pathogen defense, disease resistance, and abiotic stress response, to name a few (Provart et al., 2016). Discoveries in *Arabidopsis* (a eudicot) have been applicable to many other plant species (Piquerez et al., 2014; Provart et al., 2016) and have also impacted our understanding of human biology (Jones et al., 2008). One indication of the concentration upon, and utility of, this reference plant is that there are more papers published on *Arabidopsis* every year than there are on each of *Saccharomyces*, *Caenorhabditis*, Zebrafish, Chicken, or *Drosophila*, the predominant non-mammalian fungal and animal models (Dietrich et al., 2014). *Arabidopsis* is only one species, however, growing in a limited set of environments. *Arabidopsis*, like other members of the Brassicaceae family, has limited symbiotic relations with soil microorganisms, an annual lifestyle, C3 photosynthesis (one of several photosynthetic options), and the typical dicotyledonous ways of development. Most of our crop plants are grasses (monocots) with their own specialized physiology, while many others are legumes that fix nitrogen from the air, something *Arabidopsis* does not do. Furthermore, monocots and eudicots, such as *Arabidopsis*, are both angiosperms (flowering

\*Correspondence: carenc@umd.edu (C.C.), john.bowman@monash.edu (J.L.B.), meyerow@caltech.edu (E.M.M.).

plants) and encompass only a small fraction of the evolutionary breadth of plants. The world of plants encompasses more than 400,000 species of angiosperms, gymnosperms, ferns, lycophytes, horn-worts, mosses, liverworts, and green algae (Figure 1) representing immense biodiversity in terms of their ecosystems, architecture, biochemistry, reproductive systems, and much more. To understand the specifics of the growth of crop plants on the one hand and the diversity and evolutionary history of land plants on the other, requires study of additional species.

The properties that make a model system are of three types. The initial choice typically involves practical reasons such as economic value, or intrinsic properties that suit the species chosen for laboratory use such as small size, ease of culture, fecundity, short generation time, amenability to genetic manipulations such as crossing and mutagenesis, small genome size, and the potential to manipulate gene function. As adoption proceeds, two other types of properties are added—initially they are what might be called derived properties, such as standardized and simple procedures for DNA and RNA isolation, transformation, protein extraction, and other laboratory methods. Then, as a community of users develops, the third set of reasons to use the model are developed—community properties, such as availability of genetic strains, reporter constructs, thoroughly annotated genomes, and repositories, databases, and stock centers. Proposals to use *Arabidopsis thaliana* due to intrinsic properties date as far back as the 1940s (Laibach, 1943), and there was a seed stock center by the 1960s (Röbbelen 1965; Kranz and Kirchheim 1987). Widespread adoption of *Arabidopsis* began in the 1980s (Koornneef and Meinke, 2010; Meyerowitz and Pruitt, 1985; Somerville and Koornneef, 2002), and new seed and DNA stock centers were founded in the 1990s. The fully annotated genome (135 Mb) with 27,416 annotated genes ([http://plants.ensembl.org/Arabidopsis\\_thaliana/Info/Annotation/#assembly](http://plants.ensembl.org/Arabidopsis_thaliana/Info/Annotation/#assembly)) was complete in 2000, and by now, this species of flowering plant is central to much of the research in plant biology (Provart et al., 2016). While past model systems were chosen for practical reasons (e.g., crop plants) and/or intrinsic ones (e.g., *Arabidopsis*), the ease of genome sequencing and improved reverse genetic tools, such as TILLING (Targeting Induced Local Lesions IN Genomes), RNAi, and CRISPR/Cas9 now allow for an expansion of models that can be based on other types of considerations, such as phylogenetic relationships. Table 1 lists some of the existing and emerging plant models and reflects the current extent of sampling of taxonomic diversity among plant species.

## Algal Relatives of Land Plants

Green algae were the photosynthetic predecessors of land plants prior to the evolutionary transition to the terrestrial environment. There are two major divisions of green algae, the monophyletic chlorophytes and the paraphyletic charophytes (Figure 1), which diverged from each other well over a billion years ago (Yoon et al., 2004). Both divisions contain chlorophyll a and b, photosynthetic plastids, and cellulosic cell walls. Chlorophytes include *Chlamydomonas reinhardtii*, which is the classical model system for unicellular algae, with a long history of genetic and genomic approaches based on its intrinsic and derived properties (Harris, 2001; Merchant et al., 2007). *Chlamydomonas* can grow either photosynthetically or on acetate as a carbon source, so it is amenable to the mutational study of photosynthesis. *Chlamydomonas* bears two flagella (flagella were lost in the course of

seed plant evolution (Renzaglia and Garbary, 2001; Hodges et al., 2012), and therefore, *Chlamydomonas* has also been used extensively and productively in the study of flagellar and basal body growth and function.

Close relatives of the single-celled *Chlamydomonas* include multicellular species, an example being *Volvox carteri* (131.2 Mb; 14,247 protein-coding loci) whose bodies consist of about 2,000 somatic cells and 16 germ cells (Prochnik et al., 2010). Other members of the Volvocales, to which both *Chlamydomonas* and *Volvox* belong, exhibit intermediate levels of complexity. Comparative genomic studies within this group may provide insights into the evolution of multicellularity and organismal complexity.

With its unique features, another chlorophyte, the unicellular marine alga *Ostreococcus tauri*, has attracted interest as a “minimalist” model cell. *Ostreococcus* is the smallest known free-living eukaryotic organism and is commonly found in phytoplankton. Just 0.8 mm in diameter (Courties et al., 1994), the cell has a single nucleus, a single mitochondrion, a single Golgi body with typically 5 (up to 7) cisternae, and a single chloroplast. The ultrastructure of the cell has been determined at high resolution throughout its life cycle by cryoelectron tomography, which has revealed the cell to have just one microtubule (or two in pre-divisional cells), despite having 20 chromosomes (haploid number) to segregate in mitosis (Henderson et al., 2007). This tiny cell has only about 1,250 ribosomes (compared to 200,000 in a *Saccharomyces cerevisiae* cell) (Henderson et al., 2007). The sequenced genome is small (12.56 Mb) but has a very high gene density, with 8,166 annotated genes and intergenic regions averaging <200 bp (Derelle et al., 2006). *Ostreococcus* is easy to culture and grows rapidly. Gene analysis indicates the possibility that *Ostreococcus* has C4 photosynthesis (Derelle et al., 2006), and genome-wide metabolic network analysis has allowed a complete metabolic map (with some hypothetical reactions added) showing many additional features (Krumholz et al., 2012).

Chlorophyte algae, such as *Chlamydomonas* and *Ostreococcus*, are quite distantly related to land plants. The closest living relatives of land plants reside among the other division of green algae known as charophytes (Figure 1). Charophytes are a relatively small group of morphologically diverse freshwater and terrestrial green algae that encompass unicellular and filamentous forms, as well as structurally complex multicellular forms (Delwiche and Cooper, 2015). Land plants arose as a monophyletic group from charophyte green algae ~450 million years ago in the Ordovician Period (Delwiche and Cooper, 2015). As the closest living relatives of land plants, extant charophytes could provide insights into the origins of adaptations that allowed for the evolutionary transition to terrestrial life (Harholt et al., 2016; Delwiche and Cooper, 2015).

Despite their phylogenetic importance, charophytes have been relatively understudied and thus lack many derived and community model system properties, including axenic cultures and methods for manipulating the complete life cycle. Recent efforts, however, suggest that charophytes may be poised to become model systems, thereby filling a large phylogenetic gap between chlorophyte algae and land plants. The first charophyte draft genome sequence (104 Mb) was recently published for *Klebsormidium flaccidum* (Hori et al., 2014), a filamentous alga in one of the basal charophyte lineages. Also available are de novo

transcriptome assemblies for representative species within each of the five major charophyte lineages (Ju et al., 2015), including a desiccation response transcriptome of *Klebsormidium* (Holzinger et al., 2014). Efforts are currently underway to sequence the genomes of additional species across all the charophyte lineages. Recent studies have indicated conservation of certain plant systems in charophyte algae, including plant hormones (Holzinger and Becker, 2015; Ju et al., 2015; Viaene et al., 2013; Delaux et al., 2012), cell wall biochemistry (Domozych, 2014), and stress responses (Holzinger et al., 2014). Ethylene hormone signaling homologs in the filamentous charophyte *Spirogyra pratensis* show remarkable functional conservation with the *Arabidopsis* ethylene signaling pathway, suggesting that this hormone pathway was present in the aquatic common ancestor prior to the colonization of land (Ju et al., 2015). *Spirogyra* is within the Zygnematophyceae, a species-rich charophyte lineage that is likely to be the closest sister taxon to land plants (Delwiche and Cooper, 2015). Stable transformation has been reported for two other Zygnematophycean species (Sørensen et al., 2014; Abe et al., 2011), including the unicellular *Penium margaritaceum*. *P. margaritaceum* is being used to study the evolution of the plant cell wall, an essential adaptation for both structural support and biotic defense for life on land (Domozych, 2014).

## Land Plant Evolution

Land plants are distinguished from their algal relatives by life cycle and by morphological and physiological characteristics related to their life on land. Land plants exhibit an alternation of generations (Hofmeister, 1862), with complex multicellular bodies developing in both the haploid and diploid stages of their life cycle and, hence, are also known as embryophytes. Of the seven major extant lineages of land plants (Figure 1), three comprise the paraphyletic bryophytes (liverworts, mosses, hornworts), in which the haploid phase of the life cycle is dominant, with the diploid phase being nutritionally dependent upon the haploid. In contrast, the other four lineages (lycophytes, ferns, gymnosperms, angiosperms) comprise the monophyletic vascular plants, characterized by internal conducting vasculature and a life cycle in which the diploid phase is dominant, with the haploid phase being reduced and, in seed plants (gymnosperms and angiosperms), nutritionally dependent upon the diploid. In all land plants, except the seed plants, dispersal occurs via single-celled haploid spores and the haploid generation of the life cycle is free-living and produces gametes, motile sperm and eggs within specialized protective structures termed archegonia. Land plants are also distinguished from algae by exhibiting three-dimensional growth (as opposed to two-dimensional) and a host of biochemical adaptations, such as a waxy cuticle, associated with tolerating desiccation in a terrestrial environment.

## Bryophytes—Non-vascular Land Plants

The bryophytes (non-vascular land plants) arose at least as early as at the end of the Ordovician Period. The phylogenetic relationships both among bryophytes and with respect to vascular plants are still being debated, with almost every possible permutation having been proposed at some time. While early phylogenies built on morphology, molecules and the fossil record suggested that liverworts were the basal land plant lineage (Kenrick and Crane, 1997; Mishler and Churchill, 1984; Qiu et al., 2006). The most recent phylogeny

based on molecular data places hornworts sister to all other land plants and mosses plus liverworts as a clade sister to vascular plants (Wickett et al., 2014). However, this topology is difficult to reconcile with the fossil record, in which the earliest land plant spores (Ordovician cryptospores) more closely resemble those of extant liverworts (Wellman et al., 2003; Wellman, 2010). Likewise, the earliest Devonian bryophyte macrofossils have been classified as related to liverworts, considerably predating those thought to represent mosses and hornworts (Oostendorp, 1987; Guo et al., 2012; Edwards et al., 1995). Thus, extant liverworts appear to retain more ancestral characters of early bryophytic land plants. It is generally accepted that the three lineages of bryophytes branched from one another and from vascular plants early in the evolution of land plants, and this rapid diversification may be obscuring any remaining phylogenetic signal. Thus, to untangle early events in the origin of a land flora, models (preferably multiple) for each of the bryophyte lineages will be required.

The first bryophyte model established was the monoecious moss *Physcomitrella patens*, for which a diverse suite of techniques and resources has been developed (Knight et al., 2009). Of particular importance was the development of homologous recombination-based gene replacement, facilitating characterization of gene function (Schaefer and Zrýd, 1997). The sequence of the *Physcomitrella* genome (473 Mb; 32,926 protein-coding loci) provided an initial baseline for the identification of “land plant” genes and revealed evidence of polyploidy, a characteristic typical of many derived mosses (Rensing et al., 2008). Experiments with *Physcomitrella* have been instrumental in elucidating the molecular basis for the alternation of generations—polycomb-repressive-complex-2-mediated gene repression is required to keep the diploid gene expression program repressed during the haploid generation (Okano et al., 2009; Mosquna et al., 2009), and conversely, TALE-homeodomain genes are required to activate or maintain the diploid gene expression following gamete fusion (Sakakibara et al., 2013; Horst et al., 2016). These studies have begun to unify the molecular genetic regulation of life cycles in land plants with those in *Chlamydomonas* (Lee et al., 2008) and more distantly related eukaryotes, such as fungi (Herskowitz, 1989). While *Physcomitrella* genomics has contributed to our understanding of land plants, it is a derived moss, and to fully appreciate moss diversity, development of other model mosses (Szövényi et al., 2015b), such as *Ceratodon purpureus* (a dioecious and heavy metal-tolerant derived moss) and a *Sphagnum* species representing a basal lineage of mosses (Shaw et al., 2016), will be informative.

The dioecious liverwort *Marchantia polymorpha* has been a model for investigating a broad range of biological phenomena for almost two centuries (Bowman, 2016) and has recently entered the molecular genetic and genomic era, with all the modern tools now available (Ishizaki et al., 2016). Critically, both homologous recombination-mediated gene replacement and CRISPR/Cas9 technologies have been shown to work efficiently in *Marchantia*. Furthermore, protocols for efficient *Agrobacterium*-mediated transformation of spores have been pioneered (Ishizaki et al., 2008), and these may be able to be adapted, with modifications to suit differences in life cycle, to other spore-producing land plants. Studies on the auxin transcriptional response machinery in *M. polymorpha* have suggested that auxin response modulates, but does not determine cell fates and, furthermore, that it is critical for the transition from 2-dimensional growth to 3-dimensional growth in the haploid

generation (Kato et al., 2015; Flores-Sandoval et al., 2015). Gene knockout technologies also enabled the discovery that the molecular mechanism by which plants measure photoperiod and control the transition between vegetative growth and reproductive growth is broadly conserved across land plants (Kubota et al., 2014). These studies also hint at a lack of redundancy within the *M. polymorpha* genome (226 Mb; 19,287 protein-coding loci), a feature that would be consistent with the lack of ancient whole-genome duplications within liver-worts (Berrie, 1960).

The power of comparative genomics combined with the ability to manipulate gene function is highlighted by investigations into epidermal growth that have revealed commonalities in the development of single-celled *M. polymorpha* rhizoids, multicellular *Physcomitrella patens* rhizoids, and *Arabidopsis* root hairs. The rhizoids in the two bryophyte species develop in the haploid body while the root hairs of *Arabidopsis* are part of the other generation, the diploid. Despite these differences, orthologous basic-helix-loop-helix (bHLH) transcription factors control their development, suggesting a common ancient mechanism for epidermal outgrowths to facilitate anchorage to substrates and uptake of nutrients (Menand et al., 2007; Tam et al., 2015; Proust et al., 2016). That the *Marchantia* bHLH ortholog is also required for gemma development, also via an epidermal outgrowth, suggests that the formation of structures derived from epidermal outgrowths may have a common underlying genetic mechanism (Proust et al., 2016). As both antheridia, in which sperm will develop, and archegonia, in which egg cells are formed, are formed via epidermal outgrowths, it is possible that the formation of sex organs in most land plants will rely on a similar fundamental genetic mechanism.

Lagging behind the other bryophyte lineages in their development of models are hornworts. *Anthoceros agrestis* (83 Mb) has been proposed as a model hornwort system—culture conditions and preliminary genome sequences have been produced (Szövényi et al., 2015a). As with the other bryophyte systems, the establishment of a transformation system perhaps based on that of *M. polymorpha* would facilitate functional genetic studies. The sequencing of a hornwort genome should help to resolve the phylogenetic relationships of the three bryophytes and vascular plants by comparative genomic studies.

One caveat is that often only a single taxon in each major land plant lineage has been sequenced, as seen in Table 1. Since each of the lineages has been following their own independent evolutionary trajectory for millions of years, each species is likely to possess its own peculiarities and may not be representative of the entire lineages that they have been purported to reflect. It will be important to develop additional models within each of the lineages to conclusively address questions in evolutionary genomics.

### Basal Vascular Land Plants

Vascular plants appeared in the Silurian Period (around 423 million years ago; Kenrick and Crane, 1997). The basal lineage of vascular plants is the Lycophyta (clubmosses and relatives; Figure 1). The genome sequence of the lycophyte *Selaginella moellendorffii* (~106 Mb; 22,273 protein-coding loci) has provided some insight into the evolution of vascular plants (Banks et al., 2011); however, the development of genetic tools such as transformation methods for *Selaginella* is in the future.



The sister group to seed plants is the ferns (Figure 1), a diverse clade of plants that are often the first vascular plants to colonize disturbed sites. As with lycophytes, both haploid and diploid generations of ferns may be free-living. Ferns may be either homosporous or heterosporous and comparisons may provide insight into the evolution of sexual systems. Ferns evolved leaves independently of seed plants, and whether the molecular mechanisms are similar is actively being investigated (Plackett et al., 2015). Ferns are often polyploids with large genomes, and while no draft fern genome sequences are currently available (Wolf et al., 2015), transcriptome assemblies of *Azolla filiculoides* (an aquatic fern; genome of 740 Mb) and *Equisetum giganteum* (horsetail) have provided some insight into gene content (Vanneste et al., 2015; Brouwer et al., 2014). A derived fern, *Ceratopteris richardii* (11.26 Gb) has been developed to some extent as a model system with genetic approaches; e.g., investigations of sex determination mechanisms (Strain et al., 2001; Atallah and Banks, 2015) and transformation protocols have been established (Muthukumar et al., 2013; Plackett et al., 2014).

## Diversification in Eudicots

### *Arabidopsis* Satellite Systems

While there is a clear rationale for developing plant models that are distantly related to flowering plants, there is also justification for developing models as close as possible to the dominant flowering plant model *Arabidopsis thaliana*. These new plant models form what might be termed *Arabidopsis* satellite systems—plants that are phylogenetically as close to *Arabidopsis* as possible and therefore amenable to laboratory use and to the methods developed for *Arabidopsis* but with special features not found in *A. thaliana*. *Arabidopsis lyrata* is an example of such a satellite system. *A. lyrata* is generally self-incompatible in comparison to the generally self-crossing *A. thaliana*, so it can provide a view of genome evolution in an outcrossing species. As there are also self-fertilizing populations, the evolution of mating system differences can be studied (Joschinski et al., 2015). This species is also being used to address the evolution of genome size and chromosome number. This is an important question in plant evolution, as flowering plant genomes range in size from 82 Mb (*Utricularia gibba*; Ibarra-Laclette et al., 2013) to an estimated 17 Gb for wheat (Brenchley et al., 2012) and up to 148 Gb for the monocot *Paris japonica* (Pellicer et al., 2010). Genome sequencing of *A. lyrata* (Hu et al., 2011) shows a genome of 207 Mb, different from *A. thaliana* by thousands of small insertions/deletions, despite an overall 80% sequence similarity in comparable regions. The chromosomes show two reciprocal translocations and three chromosome fusions that led from the presumed ancestral eight chromosomes of *A. lyrata* to the five of *A. thaliana*. Half of the genome differences between the two species are transposable elements. Recent work examines gene expression after hybridization between different populations (Videvall et al., 2016). Future studies will likely include attempts to understand the basis of ecological differences between the two species, since *A. thaliana* is widely distributed in temperate climates, whereas *A. lyrata* is circumpolar and tolerant of cold (Wos and Willi, 2015).

*Arabidopsis arenosa* is another such satellite, which has both diploid and naturally formed autotetraploid populations, which has enabled the study of the evolutionary differences that

accumulate after autopolyploidy (Hollister et al., 2012), as well as differences between natural and artificially formed autotetraploid lines (Yant et al., 2013).

Another example of what could be considered an *Arabidopsis* satellite is a salt-tolerant mustard, *Thelungiella halophila* (also published under a newer designation, *Eutrema salsugineum*) (Inan et al., 2004; Yang et al., 2013). Comparative studies of osmolytes, ion channels, and responses to additional stresses are leading to an understanding of responses to abiotic stress that could be of considerable importance to agriculture in dry or drought-prone areas.

Additional *Arabidopsis* satellites are under development and are already revealing physiological and developmental processes distinct from those in *A. thaliana*. These satellites include *Cardamine hirsuta* for compound leaf development (Hay et al., 2014; Bar and Ori, 2015), *Arabidopsis halleri*, a cadmium and zinc hyperaccumulating plant (Kubota et al., 2015; Kosugi et al., 2015), *Boechera* species that exhibit apomixis (asexual reproduction) (Rushworth et al., 2011), and *Arabis alpina*, a model for the ecology of alpine and arctic environments that has a perennial rather than annual life cycle (Wang et al., 2009). Some *A. alpina* populations are self-fertilizing, while others appear to manifest sporophytic self-incompatibility (Tedder et al., 2011).

### Additional Eudicot Models

Prior to the development of *Arabidopsis* as a model system, a number of other plant models had been established, primarily among the angiosperms. Angiosperms have been a major focus of model studies as they form the largest and most diverse group of plants (Figure 1), consisting of several hundred thousand species of herbaceous plants, shrubs, grasses, and most species of trees. A variety of established and emerging models are discussed below and in subsequent sections of this article. This section focuses on plants within the asterid clade, which diverged from the rosid clade (in which *Arabidopsis* resides) early in angiosperm evolution, ~120 million years ago.

*Antirrhinum majus* (snapdragon) has numerous intrinsic qualities and derived properties, as well as an extensive history in plant research spanning the past 100 years (Schwarz-Sommer et al., 2003). The elucidation of the molecular basis of floral organ identity came from studies in *A. majus* in conjunction with studies in *Arabidopsis* (Coen and Meyerowitz, 1991). This member of the Plantaginaceae family has been also used to study the basis of floral scent, flower pigmentation, floral asymmetry, leaf development, and population genetics (Schwarz-Sommer et al., 2003). In the Solanaceae family, petunia has a similarly extensive history in plant genetics and has been recently proposed to serve as a representative model for the asterid clade (Vandenbussche et al., 2016). *Mimulus* in the Phrymaceae family of asterids is a proposed model for plant adaptation and speciation, based on the broad ecological and genomic diversity of this genus (Wu et al., 2008).

The Asteraceae or Compositae (commonly known as the daisy family) is thought to be the largest and perhaps most diverse family in the asterid clade. The most developed model within this family is lettuce (*Lactuca sativa* L.), which has a life cycle of 3–5 months, routine transformation methods, and a genome size of ~2.7 Gb (<http://compgenomics.ucdavis.edu/>)



[compositae\\_data.php?name=Lactuca+sativa](#)). Another economically valuable member of the Asteraceae is sunflower (*Helianthus annuus*), an oilseed crop that may be useful not only for oil and seed production, but also for studying plant genome evolution. The sunflower genome is quite large (~3.5 Gb) (Kane et al., 2011) mainly due to transposable elements that comprise 81% of the genome (Staton et al., 2012).

### Nitrogen Fixation

Nitrogen fixation is a physiological and developmental process that is not only fascinating, but also agriculturally of great importance. Many legumes have a symbiotic association with nitrogen fixing bacteria, whereby the plants create nitrogen-fixing nodules in their roots that provide an anaerobic environment where the bacteria can convert atmospheric nitrogen to nitrate, which can be used as a nitrogen source by the plants. Two plants used for study of nitrogen fixation are *Medicago truncatula*, an alfalfa relative with a small and sequenced genome (~454–526 Mb) (Young et al., 2011), small size, large seed production, and good transformation protocols (Barker et al., 1990; Crane et al., 2006), and *Lotus japonicus*, also with a manageable genome size (472 Mb), short life cycle, and a different (determinate versus indeterminate) type of nitrogen fixing nodule (Sato and Tabata 2006; Sato et al., 2008).

### Fruit Development

*Arabidopsis* has a dry fruit known as a silique, which is typical of the mustard family to which it belongs. *Solanum lycopersicum* (tomato) is another well-established reference plant for the study not only of fruit ripening, but also plant growth, compound leaf development, and disease resistance. There are many cultivars of tomato, and one in particular has been proposed for its *Arabidopsis*-like properties, cultivar Micro-Tom, a dwarf and determinate growth variety originally produced as an ornamental (Emmanuel and Levy, 2002; Martí et al., 2006). Its rapid life cycle (2–3 months), small size (10–20 cm), and ability to grow when planted at high density (>1,000 plants per m<sup>2</sup>) and in glasshouses suit it well for mutant screens. Insertional mutagenesis protocols allow for isolation and molecular cloning of genes that give novel mutant phenotypes, and chemical mutagenesis is also highly effective. The published genome sequences of tomato are from another, closely related cultivar and show a genome of ~900 Mb (Tomato Genome Consortium, 2012). As for *Arabidopsis*, tomatoes can be considered to have satellite systems. One of particular note is potato, a member of the same family, which serves as a model not only for tuber development (Navarro et al., 2011), but also for agriculturally relevant disease resistance.

To study the development of different types of edible fruits at a genetic and genomic level requires different models. Many fruits are non-climacteric (ripen without a burst of respiration), in contrast to tomato fruit, which is climacteric. Cultivated strawberry is non-climacteric and has a complex octaploid genome that is thought to derive from four different diploid ancestors. *Fragaria vesca*, a wild strawberry species, is diploid and has a small sequenced genome of 240 Mb (Shulaev et al., 2011). *F. vesca*, with its simpler genome, in vitro regeneration, short generation time, small size, and available transformation methods, has been proposed as a reference system for the Rosaceae family, which includes peach. *Prunus persica* (peach) also has been proposed as a model due to considerable genetic

resources, and especially because of its small genome, which at 225 Mb is less than twice that of *Arabidopsis* (International Peach Genome Initiative et al., 2013). It too is diploid, unlike sour cherries and plums, its close polyploid relatives. It cannot be said that peach is rapid-growing, however, as the time from germination to fruiting, while much shorter than in related trees, is 2 or 3 years, and its size precludes simple indoor growth.

## Monocot Models

### Grasses

Grasses constitute the most important plants for human use: the three major crops in the world are maize, rice, and wheat. Long before *Arabidopsis* came to prominence as a model, researchers studied the genetics and development of maize (*Zea mays*), as well as the genetic basis of its domestication from the wild grass teosinte, and particularly, the genetic contributions to yield and agronomic properties of this major crop (Strable and Scanlon 2009). One of the initial reasons for the use of *Arabidopsis* was the difficulty in dealing with the enormous genome of maize, which at ~2.3 Gb is similar in size to that of humans (Schnable et al., 2009). Maize is also a large and slow growing plant, with one or two generations a year, requiring considerable field space or highly illuminated growth room areas. To avoid these difficulties, and still study the growth and physiology of grasses, grass satellite systems are also intensively studied.

Rice (*Oryza sativa*) is the predominantly studied non-maize model genomic system for grasses, not only because it is a major crop itself, but also because it has a manageable genome and is a smaller plant, more easily suited to laboratory growth. The haploid genome of 389 Mb has been sequenced and annotated (International Rice Genome Sequencing Project 2005), and there are efficient transformation methods (Hiei et al., 1997). The generation time is ~4 months, depending on strain (compared to ~2 months for *Arabidopsis thaliana*).

While a great advantage of maize and rice is that anything learned is directly relevant to agriculture, there are two non-crop grasses that have been proposed as even more convenient laboratory model systems, *Brachypodium distachyon* and *Setaria viridis*. *Brachypodium* has a genome of only 272 Mb, about twice that of *Arabidopsis* and an eighth that of maize, with a small content (~20%) of repetitive elements, making cloning by chromosome walking feasible. The plant grows to 15–20 cm, similar to *Arabidopsis*, and has an 8–12 week generation time. It is transformable with high efficiency using *Agrobacterium* and tissue culture methods (Brkljacic et al., 2011). Thus, the intrinsic properties are not too different from those of *Arabidopsis*, and the derived and community aspects are under rapid development. *Brachypodium*, like rice, wheat, and *Arabidopsis*, has C3 photosynthesis, in which carbon dioxide from the air is directly combined with ribulose biphosphate to make 3-phosphoglycerate, starting the Calvin-Benson cycle. *Setaria viridis* (green foxtail) like maize uses C4 photosynthesis, in which CO<sub>2</sub> is incorporated into malate and then from malate reacts with ribulose biphosphate, allowing carbon dioxide uptake at night when humidity is high, and its use during the day without the water loss consequent on stomatal opening. *Setaria* thus can serve as a model for the specializations (anatomical and biochemical) that allow this type of light utilization, which is particularly efficient in hot and

dry conditions (Doust et al., 2009; Brutnell et al., 2010). The genome is ~395 Mbp (Bennetzen et al., 2012). An additional grass, *Oropetium thomaeum*, with the smallest known grass genome (245 Mbp), is a drought-tolerant monocot (VanBuren et al., 2015) in much the same way that *Eutrema* is a salt-tolerant satellite of *Arabidopsis*. Additional grasses are proposed and used as models for bioenergy production, discussed below.

### Monocot “Trees”

Palms (oil, date, coconut), bananas, and pineapples have great economic importance in arid, tropical, and/or subtropical regions. These monocotyledonous flowering plants are sometimes referred to as “trees,” although they are distinct from woody plants that have secondary xylem growth. Comparison of the genome sequences of diploid African oil palm (*Elaeis guineensis*; 1.8 Gb) and diploid South American oil palm (*E. olifera*) revealed that their common ancestor was tetraploid (Singh et al., 2013). Genome sequencing is underway or nearly complete for date palm (*Phoenix dactylifera*; 658 Mb) (Al-Dous et al., 2011), banana (*Musa acuminata*; 523 Mb) (D'Hont et al., 2012), and pineapple (*Ananas comosus*; 526 Mb) (Ming et al., 2015).

### Bioenergy Crops

An additional important use for plants beyond their use for food is their use for energy, either by burning, the traditional route, or for conversion to high-energy transportation fuels (bioethanol and biodiesel). *Sorghum bicolor* is one model energy crop, with low input, high drought tolerance, and accumulation of large amounts of sugars in stems, which grow to high biomass (Carpita and McCann 2008; Calviño and Messing 2012). It is also a potential model for study of C4 photosynthesis (Mullet et al., 2014). The ~732 Mb genome is partly sequenced (Paterson et al., 2009). Genomic analysis by high-resolution genetic mapping and sequencing of the sugarcane relative *Miscanthus sinensis* (a parental species of the hybrid energy crop *Miscanthus x giganteus*) has also been started (Ma et al., 2012; Swaminathan et al., 2012) as has sequencing of *Miscanthus x giganteus* (Swaminathan et al., 2010). Another high-yielding energy crop is switch-grass, *Panicum virgatum*, which is also the subject of genome analysis projects that show a close relationship to sorghum (Sharma et al., 2012). The duckweeds are highly reduced aquatic monocots. A truly tiny (2 cm) member, the fast-growing *Lemna minor*, has additionally been proposed as a model for aquatic plants, as a biofuel source, and because of its sensitivity to environmental toxins, for ecophysiological studies (Van Hoeck et al., 2015). The diploid genome is 472 Mb (Van Hoeck et al., 2015), and efficient genetic transformation and gene silencing via artificial microRNAs have been established (Cantó-Pastor et al., 2015). The smallest known monocot genome (158 Mb) belongs to another duckweed, *Spirodela polyrrhiza*, whose genome sequence suggests that reductions in several biochemical pathways accompany its aquatic lifestyle and that its neotenus life cycle is reflected in changes in paralog numbers of promoters and repressors of the juvenile-to-adult transition (Wang et al., 2014).

### Medicinal Plants

There are hundreds of medicinal plants, and many of our drugs are either derived from plants or based on compounds originally identified in plants, e.g., aspirin from *Salix* species,

morphine from *Papaver somniferum*, digitalin from *Digitalis purpurea*, quinine from *Cinchona* species, vincristine from *Catharanthus roseus*, and taxol from *Taxus* species. However, the compounds often exist in minute quantities, in sometimes rare plants, making the identification of the genetic basis for biochemical pathways and their transfer to more tractable systems of both economic and environmental interest. Targeted genomic or transcriptomic sequencing of medicinal plant species (Hao et al., 2012; Sharma and Shrivastava, 2016) could lead to synthetic biology approaches to produce specific bioactive compounds in yeast or *E. coli*, as has been done with *Artemisia annua*, the source of artemisinin, a potent anti-malarial drug (Ro et al., 2006; Westfall et al., 2012), and the opium poppy *Papaver somniferum*, the source of medical opiates (Thodey et al., 2014).

## Trees

Trees (both eudicot hardwoods and gymnosperm softwoods) exhibit dormancy and secondary wood formation, and have economic, environmental, and ecological importance, due to their use for wood-based products (timber, pulp, and paper), carbon mitigation, and watershed protection. The reference tree *Populus trichocarpa* (black cottonwood or California poplar) was the first tree (and the third angiosperm, after *Arabidopsis* and then rice) to have its genome (~422.9 Mb) sequenced (Tuskan et al., 2006) and is readily transformable (Song et al., 2006). The genus *Populus* includes poplars, cottonwoods, and aspens. Poplar is a valuable source of timber because of its rapid growth, maturing in ~5 years. Genome information from poplar has been used to study wood formation, salt stress, and other relevant characteristics (Dharmawardhana et al., 2010; Wullschleger et al., 2013; Zang et al., 2015). The genome sequences of the fast-growing hardwood trees, *Eucalyptus grandis* (640 Mb) and *E. globulus* (530 Mb), offer the opportunity for comparative angiosperm tree genomics as well as investigation of specialized metabolites, such as terpenes and oils (Myburg et al., 2014). While not a model for trees per se, transcriptome sequencing of *Ulmus minor* (elm) genotypes has shed insight on the impact that Dutch elm disease has had on *Ulmus* species and may illuminate possible avenues for conservation of *Ulmus* species (Perdoguero et al., 2015).

Gymnosperms, the sister group of angiosperms, are the oldest lineage of the seed plants (Figure 1). Gymnosperms are mostly long-lived species with relatively large genomes, so there is a dearth of model gymnosperms at even the first level, despite genome sequences being available for a few northern hemisphere conifers: the Norway spruce (20 Gb genome) (*Picea abies*; Nystedt et al., 2013), white spruce (20 Gb) (*Picea glauca*; Birol et al., 2013), and the loblolly pine (22 Gb) (*Pinus taeda*; Zimin et al., 2014). However, due to their value as timber, genome-wide association mapping approaches have been applied in conifer species, identifying allelic variation influencing wood properties and environmental responses (González-Martínez et al., 2007; Dillon et al., 2010; Eckert et al., 2010).

## Conclusion

In summary, both the list of what are considered model systems for plants and the definition of model systems, are undergoing rapid change. Few of the plants listed have the full set of intrinsic, derived, and community properties that characterize the established models

*Arabidopsis*, tomato, maize, rice, and *Chlamydomonas*. What they do have is a group of laboratories that have particular reasons, based on unique physiology, development, phylogenetic position, or societal value to warrant deeper study, and a genome sequence (or a plan to obtain one) that can give the list of parts—important prerequisites for a generally accepted model status. Many also have demonstrated transformation protocols, amenable genetic systems, and ease of laboratory culture. Classical genetic manipulation, which was a critical feature of *Arabidopsis*, may no longer be essential, given the availability of rapid whole-genome sequencing and targeted gene editing by CRISPR/Cas9. Whether the future holds continued concentration on the current set of models, a more diverse approach in which the overall effort of the plant research community is spread among many systems—a “return” to the situation prior to the advent of *Arabidopsis* as a model, a greater concentration on crop plants, or some mix of the above, is yet to be seen. One thing that is clear is that the choice of possibilities for plant researchers is large and open—and that any information obtained about any plant model will be applicable to studies of other plants, improving our understanding of the evolution, diversification, and fundamental properties of plants. The expansion of knowledge will enhance our ability to modify plants and make use of plant biodiversity. Like the original discoveries in plants of cells, the nucleus, genetic principles, genes, viruses, transposons, the cytoskeleton, microRNAs, DNA methylation, and much else, the information may also be relevant to those who study animals.

## ACKNOWLEDGMENTS

The laboratory of C.C. is supported by NSF grant MCB-1244303. C.C. is supported in part by the Maryland Agricultural Experiment Station. The laboratory of J.L.B. is supported by the Australian Research Council (DP160100892). The laboratory of E.M.M. is supported by funds from the Howard Hughes Medical Institute, the Gordon and Betty Moore Foundation (through Grant GBMF3406), and the NIH (through Grant R01 GM104244).

## REFERENCES

- Abe J, Hori S, Tsuchikane Y, Kitao N, Kato M, Sekimoto H. Stable nuclear transformation of the *Closterium peracerosum-strigosum-littorale* complex. *Plant Cell Physiol.* 2011; 52:1676–1685. [PubMed: 21804095]
- Al-Dous EK, George B, Al-Mahmoud ME, Al-Jaber MY, Wang H, Salameh YM, Al-Azwani EK, Chaluvadi S, Pontaroli AC, DeBarry J, et al. De novo genome sequencing and comparative genomics of date palm (*Phoenix dactylifera*). *Nat. Biotechnol.* 2011; 29:521–527. [PubMed: 21623354]
- Amborella Genome Project. The Amborella genome and the evolution of flowering plants. *Science.* 2013; 342:1241089. [PubMed: 24357323]
- Araújo P, Cesarino I, Mayer JL, Ferrari IF, Kiyota E, Sawaya AC, Paes Leme AF, Mazzafera P. A model system to study the lignification process in *Eucalyptus globulus*. *Physiol. Plant.* 2014; 152:17–31. [PubMed: 24444279]
- Atallah NM, Banks JA. Reproduction and the pheromonal regulation of sex type in fern gametophytes. *Front. Plant Sci.* 2015; 6:100. [PubMed: 25798139]
- Banks JA, Nishiyama T, Hasebe M, Bowman JL, Gribskov M, dePamphilis C, Albert VA, Aono N, Aoyama T, Ambrose BA, et al. The *Selaginella* genome identifies genetic changes associated with the evolution of vascular plants. *Science.* 2011; 332:960–963. [PubMed: 21551031]
- Bar M, Ori N. Compound leaf development in model plant species. *Curr. Opin. Plant Biol.* 2015; 23:61–69. [PubMed: 25449728]
- Barker DG, Bianchi S, Blondon F, Dattée Y, Duc G, Essad S, Flament P, Gallusci P, Génier G, Guy P, et al. *Medicago truncatula*, a model plant for studying the molecular genetics of the Rhizobium-legume symbiosis. *Plant Mol. Biol. Report.* 1990; 8:40–49.

- Bennetzen JL, Schmutz J, Wang H, Percifield R, Hawkins J, Pontaroli AC, Estep M, Feng L, Vaughn JN, Grimwood J, et al. Reference genome sequence of the model plant *Setaria*. *Nat. Biotechnol.* 2012; 30:555–561. [PubMed: 22580951]
- Berrie GK. The chromosome numbers of liverworts (Hepaticae and Anthocerotae). *Trans. Brit. Bryol. Soc.* 1960; 3:688–705.
- Birol I, Raymond A, Jackman SD, Pleasance S, Coope R, Taylor GA, Yuen MMS, Keeling CI, Brand D, Vandervalk BP, et al. Assembling the 20 Gb white spruce (*Picea glauca*) genome from whole-genome shotgun sequencing data. *Bioinformatics.* 2013; 29:1492–1497. [PubMed: 23698863]
- Bowman JL. A brief history of *Marchantia* from Greece to genomics. *Plant Cell Physiol.* 2016; 57:210–229. [PubMed: 25766905]
- Bowman JL, Floyd SK, Sakakibara K. Green genes-comparative genomics of the green branch of life. *Cell.* 2007; 129:229–234. [PubMed: 17448980]
- Brenchley R, Spannagl M, Pfeifer M, Barker GLA, D'Amore R, Allen AM, McKenzie N, Kramer M, Kerhornou A, Bolser D, et al. Analysis of the bread wheat genome using whole-genome shotgun sequencing. *Nature.* 2012; 491:705–710. [PubMed: 23192148]
- Brlkjacic J, Grotewold E, Scholl R, Mockler T, Garvin DF, Vain P, Brut-nell T, Sibout R, Bevan M, Budak H, et al. Brachypodium as a model for the grasses: today and the future. *Plant Physiol.* 2011; 157:3–13. [PubMed: 21771916]
- Brouwer P, Bräutigam A, Kūlahoglu C, Tazelaar AOE, Kurz S, Nierop KGJ, van der Werf A, Weber APM, Schluepmann H. *Azolla* domestication towards a biobased economy? *New Phytol.* 2014; 202:1069–1082. [PubMed: 24494738]
- Brutnell TP, Wang L, Swartwood K, Goldschmidt A, Jackson D, Zhu X-G, Kellogg E, Van Eck J. *Setaria viridis*: a model for C<sub>4</sub> photosynthesis. *Plant Cell.* 2010; 22:2537–2544. [PubMed: 20693355]
- Cai J, Liu X, Vanneste K, Proost S, Tsai W-C, Liu K-W, Chen L-J, He Y, Xu Q, Bian C, et al. The genome sequence of the orchid *Phalaenopsis equestris*. *Nat. Genet.* 2015; 47:65–72. [PubMed: 25420146]
- Calviño M, Messing J. Sweet sorghum as a model system for bioenergy crops. *Curr. Opin. Biotechnol.* 2012; 23:323–329. [PubMed: 22204822]
- Cantó-Pastor A, Mollá-Morales A, Ernst E, Dahl W, Zhai J, Yan Y, Meyers BC, Shanklin J, Martienssen R. Efficient transformation and artificial miRNA gene silencing in *Lemna minor*. *Plant Biol (Stuttg).* 2015; 17(Suppl 1):59–65. [PubMed: 24989135]
- Carpita NC, McCann MC. Maize and sorghum: genetic resources for bioenergy grasses. *Trends Plant Sci.* 2008; 13:415–420. [PubMed: 18650120]
- Cichorz S, Goska M, Rewers M. *Miscanthus*: Inter- and intraspecific genome size variation among *M. × Giganteus*, *M. Sinensis*, *M. Sacchariflorus* accessions. *Acta Biol. Cracov. Ser.; Bot.* 2015; 57:104–133.
- Coen ES, Meyerowitz EM. The war of the whorls: genetic interactions controlling flower development. *Nature.* 1991; 353:31–37. [PubMed: 1715520]
- Courties C, Vaquer A, Troussellier M, Lautier J, Chrétiennot-Dinet MJ, Neveux J, Machado C, Claustre H. Smallest eukaryotic organism. *Nature.* 1994; 370:255.
- Crane C, Wright E, Dixon RA, Wang ZY. Transgenic *Medicago truncatula* plants obtained from *Agrobacterium tumefaciens* -transformed roots and *Agrobacterium rhizogenes*-transformed hairy roots. *Planta.* 2006; 223:1344–1354. [PubMed: 16575594]
- D'Hont A, Denoeud F, Aury J-M, Baurens F-C, Carreel F, Garsmeur O, Noel B, Bocs S, Droc G, Rouard M, et al. The banana (*Musa acuminata*) genome and the evolution of monocotyledonous plants. *Nature.* 2012; 488:213–217. [PubMed: 22801500]
- Delaux P-M, Xie X, Timme RE, Puech-Pages V, Dunand C, Lecompte E, Delwiche CF, Yoneyama K, Bécard G, Séjalon-Delmas N. Origin of strigolactones in the green lineage. *New Phytol.* 2012; 195:857–871. [PubMed: 22738134]
- Delwiche CF, Cooper ED. The evolutionary origin of a terrestrial flora. *Curr. Biol.* 2015; 25:R899–R910. [PubMed: 26439353]
- Derelle E, Ferraz C, Rombauts S, Rouzé P, Worden AZ, Robbens S, Partensky F, Degroeve S, Echeynié S, Cooke R, et al. Genome analysis of the smallest free-living eukaryote *Ostreococcus*



*tauri* unveils many unique features. Proc. Natl. Acad. Sci. USA. 2006; 103:11647–11652. [PubMed: 16868079]

- Dharmawardhana P, Brunner AM, Strauss SH. Genome-wide transcriptome analysis of the transition from primary to secondary stem development in *Populus trichocarpa*. BMC Genomics. 2010; 11:150. [PubMed: 20199690]
- Dietrich MR, Ankeny RA, Chen PM. Publication trends in model organism research. Genetics. 2014; 198:787–794. [PubMed: 25381363]
- Dillon SK, Nolan M, Li W, Bell C, Wu HX, Southerton SG. Allelic variation in cell wall candidate genes affecting solid wood properties in natural populations and land races of *Pinus radiata*. Genetics. 2010; 185:1477–1487. [PubMed: 20498299]
- Domozych DS. *Penium margaritaceum*: A unicellular model organism for studying plant cell wall architecture and dynamics. Plants (Basel). 2014; 3:543–558. [PubMed: 27135519]
- Doust AN, Kellogg EA, Devos KM, Bennetzen JL. Foxtail millet: a sequence-driven grass model system. Plant Physiol. 2009; 149:137–141. [PubMed: 19126705]
- Eckert AJ, van Heerwaarden J, Wegrzyn JL, Nelson CD, Ross-Ibarra J, González-Martínez SC, Neale DB. Patterns of population structure and environmental associations to aridity across the range of loblolly pine (*Pinus taeda* L., Pinaceae). Genetics. 2010; 185:969–982. [PubMed: 20439779]
- Edwards D, Duckett JG, Richardson JB. Hepatic characters in the earliest land plants. Nature. 1995; 374:635–636.
- Emmanuel E, Levy AA. Tomato mutants as tools for functional genomics. Curr. Opin. Plant Biol. 2002; 5:112–117. [PubMed: 11856605]
- Flores-Sandoval E, Eklund DM, Bowman JL. A simple auxin transcriptional response system regulates multiple morphogenetic processes in the liverwort *Marchantia polymorpha*. PLoS Genet. 2015; 11:e1005207. [PubMed: 26020649]
- González-Martínez SC, Wheeler NC, Ersoz E, Nelson CD, Neale DB. Association genetics in *Pinus taeda* L. I. Wood property traits. Genetics. 2007; 175:399–409. [PubMed: 17110498]
- Guo CQ, Edwards D, Wu PC, Duckett JG, Hueber FM, Li C-S. *Riccardiathallus devonicus* gen. et sp. nov., the earliest simple thalloid liverwort from the Lower Devonian of Yunnan, China. Rev. Palaeobot. Palynol. 2012; 176-177:35–40.
- Hao D-C, Chen S-L, Xiao P-G, Liu M. Application of high-throughput sequencing in medicinal plant transcriptome studies. Drug Development Research. 2012; 73:487–498.
- Harholt J, Moestrup Ø, Ulvskov P. Why plants were terrestrial from the beginning. Trends Plant Sci. 2016; 21:96–101. [PubMed: 26706443]
- Harris EH. Chlamydomonas as a model organism. Annu. Rev. Plant Physiol. Plant Mol. Biol. 2001; 52:363–406. [PubMed: 11337403]
- Hay AS, Pieper B, Cooke E, Mandáková T, Cartolano M, Tattersall AD, Ioio RD, McGowan SJ, Barkoulas M, Galinha C, et al. Cardamine hirsuta: a versatile genetic system for comparative studies. Plant J. 2014; 78:1–15. [PubMed: 24460550]
- Henderson GP, Gan L, Jensen GJ. 3-D ultrastructure of *O. tauri*: electron cryotomography of an entire eukaryotic cell. PLoS ONE. 2007; 2:e749. [PubMed: 17710148]
- Herskowitz I. A regulatory hierarchy for cell specialization in yeast. Nature. 1989; 342:749–757. [PubMed: 2513489]
- Hiei Y, Komari T, Kubo T. Transformation of rice mediated by *Agrobacterium tumefaciens*. Plant Mol. Biol. 1997; 35:205–218. [PubMed: 9291974]
- Hodges ME, Wickstead B, Gull K, Langdale JA. The evolution of land plant cilia. New Phytol. 2012; 195:526–540. [PubMed: 22691130]
- Hofmeister, W. On the Germination, Development, and Fructification of the Higher Cryptogamia and on the Fructification of the Coniferae. Currey, F., editor. Ray Society; London: 1862.
- Hollister JD, Arnold BJ, Svedin E, Xue KS, Dilkes BP, Bomblies K. Genetic adaptation associated with genome-doubling in autotetraploid *Arabidopsis arenosa*. PLoS Genet. 2012; 8:e1003093. [PubMed: 23284289]
- Holzinger A, Becker B. Desiccation tolerance in the streptophyte green alga *Klebsormidium*: The role of phytohormones. Commun. Integr. Biol. 2015; 8:e1059978. [PubMed: 26479261]

- Holzinger A, Kaplan F, Blaas K, Zechmann B, Komsic-Buchmann K, Becker B. Transcriptomics of desiccation tolerance in the streptophyte green alga *Klebsormidium* reveal a land plant-like defense reaction. PLoS ONE. 2014; 9:e110630. [PubMed: 25340847]
- Hori K, Maruyama F, Fujisawa T, Togashi T, Yamamoto N, Seo M, Sato S, Yamada T, Mori H, Tajima N, et al. *Klebsormidium flaccidum* genome reveals primary factors for plant terrestrial adaptation. Nat. Commun. 2014; 5:3978. [PubMed: 24865297]
- Horst NA, Katz A, Pereman I, Decker EL, Ohad N, Reski R. A single homeobox gene triggers phase transition, embryogenesis and asexual reproduction. Nat Plants. 2016; 2:15209. [PubMed: 27250874]
- Hsu CC, Chung YL, Chen TC, Lee YL, Kuo YT, Tsai WC, Hsiao YY, Chen YW, Wu WL, Chen HH. An overview of the Phalaenopsis orchid genome through BAC end sequence analysis. BMC Plant Biol. 2011; 11:3. [PubMed: 21208460]
- Hu TT, Pattyn P, Bakker EG, Cao J, Cheng JF, Clark RM, Fahlgren N, Fawcett JA, Grimwood J, Gundlach H, et al. The *Arabidopsis lyrata* genome sequence and the basis of rapid genome size change. Nat. Genet. 2011; 43:476–481. [PubMed: 21478890]
- Ibarra-Laclette E, Lyons E, Hernández-Guzmán G, Pérez-Torres CA, Carretero-Paulet L, Chang TH, Lan T, Welch AJ, Juárez MJ, Simpson J, et al. Architecture and evolution of a minute plant genome. Nature. 2013; 498:94–98. [PubMed: 23665961]
- Inan G, Zhang Q, Li P, Wang Z, Cao Z, Zhang H, Zhang C, Quist TM, Goodwin SM, Zhu J, et al. Salt cress. A halophyte and cryophyte Arabidopsis relative model system and its applicability to molecular genetic analyses of growth and development of extremophiles. Plant Physiol. 2004; 135:1718–1737. [PubMed: 15247369]
- International Rice Genome Sequencing Project. The map-based sequence of the rice genome. Nature. 2005; 436:793–800. [PubMed: 16100779]
- Ishizaki K, Chiyoda S, Yamato KT, Kohchi T. Agrobacterium-mediated transformation of the haploid liverwort *Marchantia polymorpha* L., an emerging model for plant biology. Plant Cell Physiol. 2008; 49:1084–1091. [PubMed: 18535011]
- Ishizaki K, Nishihama R, Yamato KT, Kohchi T. Molecular genetic tools and techniques for *Marchantia polymorpha* research. Plant Cell Physiol. 2016; 57:262–270. [PubMed: 26116421]
- Jaillon O, Aury J-M, Noel B, Policriti A, Clepet C, Casagrande A, Choisne N, Aubourg S, Vitulo N, Jubin C, et al. French-Italian Public Consortium for Grapevine Genome Characterization. The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. Nature. 2007; 449:463–467. [PubMed: 17721507]
- Jones AM, Chory J, Dangel JL, Estelle M, Jacobsen SE, Meyerowitz EM, Nordborg M, Weigel D. The impact of *Arabidopsis* on human health: diversifying our portfolio. Cell. 2008; 133:939–943. [PubMed: 18555767]
- Joschinski J, van Kleunen M, Stift M. Costs associated with the evolution of selfing in North American populations of *Arabidopsis lyrata*? Evol. Ecol. 2015; 29:749–764.
- Ju C, Van de Poel B, Cooper ED, Thierer JH, Gibbons TR, Delwiche CF, Chang C. Conservation of ethylene as a plant hormone over 450 million years of evolution. Nat Plants. 2015; 1:14004. [PubMed: 27246051]
- Kane NC, Gill N, King MG, Bowers JE, Berges H, Gouzy J, Bachlava E, Langlade NB, Lai Z, Stewart M, et al. Progress towards a reference genome for sunflower. Botany. 2011; 89:429–437.
- Kato H, Ishizaki K, Kouno M, Shirakawa M, Bowman JL, Nishihama R, Kohchi T. Auxin-mediated transcriptional system with a minimal set of components is critical for morphogenesis through the life cycle in *Marchantia polymorpha*. PLoS Genet. 2015; 11:e1005084. [PubMed: 26020919]
- Kenrick, P.; Crane, PR. The origin and early diversification of land plants: a cladistic study. Smithsonian Institution Press; Washington: 1997.
- Klebs, G. Die Bedingungen der Fortpflanzung bei einigen Algen und Pilzen. Gustav Fischer; Jena: 1896.
- Knight, C.; Perroud, P-F.; Cove, D. The moss *Physcomitrella patens*, Ann. Plant Rev. Vol. 36. Wiley-Blackwell; 2009.
- Koornneef M, Meinke D. The development of *Arabidopsis* as a model plant. Plant J. 2010; 61:909–921. [PubMed: 20409266]

- Kosugi A, Tamaru J, Gotou K, Furihata HY, Shinuzu A, Kawabe A, Harada E. Metal accumulation by *Arabidopsis halleri* subsp. *gemmifera* at a limestone mining site. *Aust. J. Bot.* 2015; 63:134.
- Kramer EM, Hodges SA. Aquilegia as a model system for the evolution and ecology of petals. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 2010; 365:477–490. [PubMed: 20047874]
- Kranz, AR.; Kirchheim, B. Genetic Resources in *Arabidopsis*. Arabidopsis Information Service; 1987.
- Krumholz EW, Yang H, Weisenhorn P, Henry CS, Libourel IG. Genome-wide metabolic network reconstruction of the picoalga *Ostreococcus*. *J. Exp. Bot.* 2012; 63:2353–2362. [PubMed: 22207618]
- Kubota A, Kita S, Ishizaki K, Nishihama R, Yamato KT, Kohchi T. Co-option of a photoperiodic growth-phase transition system during land plant evolution. *Nat. Commun.* 2014; 5:3668. [PubMed: 24752248]
- Kubota S, Iwasaki T, Hanada K, Nagano AJ, Fujiyama A, Toyoda A, Sugano S, Suzuki Y, Hikosaka K, Ito M, Morinaga S. A genome scan for genes underlying microgeographic-scale local adaptation in a wild *Arabidopsis* species. *PLoS Genet.* 2015; 11:e1005361. [PubMed: 26172569]
- Laibach F. *Arabidopsis thaliana* (L.) Heynh. als Objekt für genetische und entwicklungsphysiologische Untersuchungen. *Bot. Arch.* 1943; 44:439–455.
- Lauterborn R. Zeitschrift für wissenschaftliche, Protozoen-Studien. IV. Theil. Flagellaten aus dem Gebiete des Oberrheins. *Zoologie.* 1898; 65:369–391.
- Lee J-H, Lin H, Joo S, Goodenough U. Early sexual origins of homeoprotein heterodimerization and evolution of the plant KNOX/BELL family. *Cell.* 2008; 133:829–840. [PubMed: 18510927]
- Ma X-F, Jensen E, Alexandrov N, Troukhan M, Zhang L, Thomas-Jones S, Farrar K, Clifton-Brown J, Donnison I, Swaller T, Flavell R. High resolution genetic mapping by genome sequencing reveals genome duplication and tetraploid genetic structure of the diploid *Miscanthus sinensis*. *PLoS ONE.* 2012; 7:e33821. [PubMed: 22439001]
- Martí E, Gisbert C, Bishop GJ, Dixon MS, García-Martínez JL. Genetic and physiological characterization of tomato cv. Micro-Tom. *J. Exp. Bot.* 2006; 57:2037–2047. [PubMed: 16687436]
- Menand B, Yi K, Jouannic S, Hoffmann L, Ryan E, Linstead P, Schaefer DG, Dolan L. An ancient mechanism controls the development of cells with a rooting function in land plants. *Science.* 2007; 316:1477–1480. [PubMed: 17556585]
- Merchant SS, Prochnik SE, Vallon O, Harris EH, Karpowicz SJ, Witman GB, Terry A, Salamov A, Fritz-Laylin LK, Maréchal-Drouard L, et al. The *Chlamydomonas* genome reveals the evolution of key animal and plant functions. *Science.* 2007; 318:245–250. [PubMed: 17932292]
- Meyerowitz EM, Pruitt RE. *Arabidopsis thaliana* and plant molecular genetics. *Science.* 1985; 229:1214–1218. [PubMed: 17770799]
- Ming R, VanBuren R, Wai CM, Tang H, Schatz MC, Bowers JE, Lyons E, Wang ML, Chen J, Biggers E, et al. The pineapple genome and the evolution of CAM photosynthesis. *Nat. Genet.* 2015; 47:1435–1442. [PubMed: 26523774]
- Mishler BD, Churchill SP. A cladistic approach to the phylogeny of the bryophytes. *Brittonia.* 1984; 36:406–424.
- Mosquna A, Katz A, Decker EL, Rensing SA, Reski R, Ohad N. Regulation of stem cell maintenance by the Polycomb protein FIE has been conserved during land plant evolution. *Development.* 2009; 136:2433–2444. [PubMed: 19542356]
- Mullet J, Morishige D, McCormick R, Truong S, Hilley J, McKinley B, Anderson R, Olson SN, Rooney W. Energy sorghum—a genetic model for the design of C4 grass bioenergy crops. *J. Exp. Bot.* 2014; 65:3479–3489. [PubMed: 24958898]
- Muthukumar B, Joyce BL, Elless MP, Stewart CN Jr. Stable transformation of ferns using spores as targets: *Pteris vittata* and *Ceratopteris thalictroides*. *Plant Physiol.* 2013; 163:648–658. [PubMed: 23933990]
- Myburg AA, Grattapaglia D, Tuskan GA, Hellsten U, Hayes RD, Grimwood J, Jenkins J, Lindquist E, Tice H, Bauer D, et al. The genome of *Eucalyptus grandis*. *Nature.* 2014; 510:356–362. [PubMed: 24919147]
- Navarro C, Abelenda JA, Cruz-Oró E, Cuéllar CA, Tamaki S, Silva J, Shimamoto K, Prat S. Control of flowering and storage organ formation in potato by FLOWERING LOCUS T. *Nature.* 2011; 478:119–122. [PubMed: 21947007]

- Nystedt B, Street NR, Wetterbom A, Zuccolo A, Lin YC, Scofield DG, Vezzi F, Delhomme N, Giacomello S, Alexeyenko A, et al. The Norway spruce genome sequence and conifer genome evolution. *Nature*. 2013; 497:579–584. [PubMed: 23698360]
- Okano Y, Aono N, Hiwatashi Y, Murata T, Nishiyama T, Ishikawa T, Kubo M, Hasebe M. A polycomb repressive complex 2 gene regulates apogamy and gives evolutionary insights into early land plant evolution. *Proc. Natl. Acad. Sci. USA*. 2009; 106:16321–16326. [PubMed: 19805300]
- Oostendorp C. The Bryophytes of the Paleozoic and Mesozoic. *Bryophyt. Bibl*. 1987; 34:1–110.
- Pan Z-J, Chen Y-Y, Du J-S, Chen Y-Y, Chung M-C, Tsai W-C, Wang CN, Chen H-H. Flower development of *Phalaenopsis* orchid involves functionally divergent *SEPALLATA*-like genes. *New Phytol*. 2014; 202:1024–1042. [PubMed: 24571782]
- Paterson AH, Bowers JE, Bruggmann R, Dubchak I, Grimwood J, Gundlach H, Haberer G, Hellsten U, Mitros T, Poliakov A, et al. The *Sorghum bicolor* genome and the diversification of grasses. *Nature*. 2009; 457:551–556. [PubMed: 19189423]
- Pellicer J, Fay MF, Leitch IJ. The largest eukaryotic genome of them all? *Bot. J. Linn. Soc*. 2010; 164:10–15.
- Perdoguero P, Venturas M, Cervera MT, Gil L, Collada C. Massive sequencing of *Ulmus minor*'s transcriptome provides new molecular tools for a genus under the constant threat of Dutch elm disease. *Front. Plant Sci*. 2015; 6:541. [PubMed: 26257751]
- Piquerez SJM, Harvey SE, Beynon JL, Ntoukakis V. Improving crop disease resistance: lessons from research on *Arabidopsis* and tomato. *Front. Plant Sci*. 2014; 5:671. [PubMed: 25520730]
- Plackett ARG, Huang L, Sanders HL, Langdale JA. High-efficiency stable transformation of the model fern species *Ceratopteris richardii* via microparticle bombardment. *Plant Physiol*. 2014; 165:3–14. [PubMed: 24623851]
- Plackett ARG, Di Stilio VS, Langdale JA. Ferns: the missing link in shoot evolution and development. *Front. Plant Sci*. 2015; 6:972. [PubMed: 26594222]
- Prochnik SE, Umen J, Nedelcu AM, Hallmann A, Miller SM, Nishii I, Ferris P, Kuo A, Mitros T, Fritz-Laylin LK, et al. Genomic analysis of organismal complexity in the multicellular green alga *Volvox carteri*. *Science*. 2010; 329:223–226. [PubMed: 20616280]
- Proust H, Honkanen S, Jones VA, Morieri G, Prescott H, Kelly S, Ishizaki K, Kohchi T, Dolan L. RSL class I genes controlled the development of epidermal structures in the common ancestor of land plants. *Curr. Biol*. 2016; 26:93–99. [PubMed: 26725198]
- Provart NJ, Alonso J, Assmann SM, Bergmann D, Brady SM, Brkljacic J, Browse J, Chapple C, Colot V, Cutler S, et al. 50 years of *Arabidopsis* research: highlights and future directions. *New Phytol*. 2016; 209:921–944. [PubMed: 26465351]
- Pulvirenti A, Giugno R, Distefano R, Pigola G, Mongiovi M, Giudice G, Vendramin V, Lombardo A, Cattonaro F, Ferro A. A knowledge base for *Vitis vinifera* functional analysis. *BMC Syst. Biol*. 2015; 9(Suppl 3):S5.
- Qiu Y-L, Yu J. *Azolla*—a model organism for plant genomic studies. *Genomics Proteomics Bioinformatics*. 2003; 1:15–25. [PubMed: 15626330]
- Qiu YL, Li L, Wang B, Chen Z, Knoop V, Groth-Malonek M, Dombrowska O, Lee J, Kent L, Rest J, et al. The deepest divergences in land plants inferred from phylogenomic evidence. *Proc. Natl. Acad. Sci. USA*. 2006; 103:15511–15516. [PubMed: 17030812]
- Rayburn AL, Crawford J, Rayburn CM, Juvik JA. Genome size of three *Miscanthus* species. *Plant Mol. Biol. Report*. 2009; 27:184–188.
- Rensing SA, Lang D, Zimmer AD, Terry A, Salamov A, Shapiro H, Nishiyama T, Perroud PF, Lindquist EA, Kamisugi Y, et al. The *Physcomitrella* genome reveals evolutionary insights into the conquest of land by plants. *Science*. 2008; 319:64–69. [PubMed: 18079367]
- Renzaglia KS, Garbary DJ. Motile gametes of land plants: diversity, development, and evolution. *Crit. Rev. Plant Sci*. 2001; 20:107–213.
- Ro D-K, Paradise EM, Ouellet M, Fisher KJ, Newman KL, Ndungu JM, Ho KA, Eachus RA, Ham TS, Kirby J, et al. Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature*. 2006; 440:940–943. [PubMed: 16612385]
- Röbbelen G. The Laibach standard collection of natural races. *Arabidopsis Information Service*. 1965; 2:36–47.

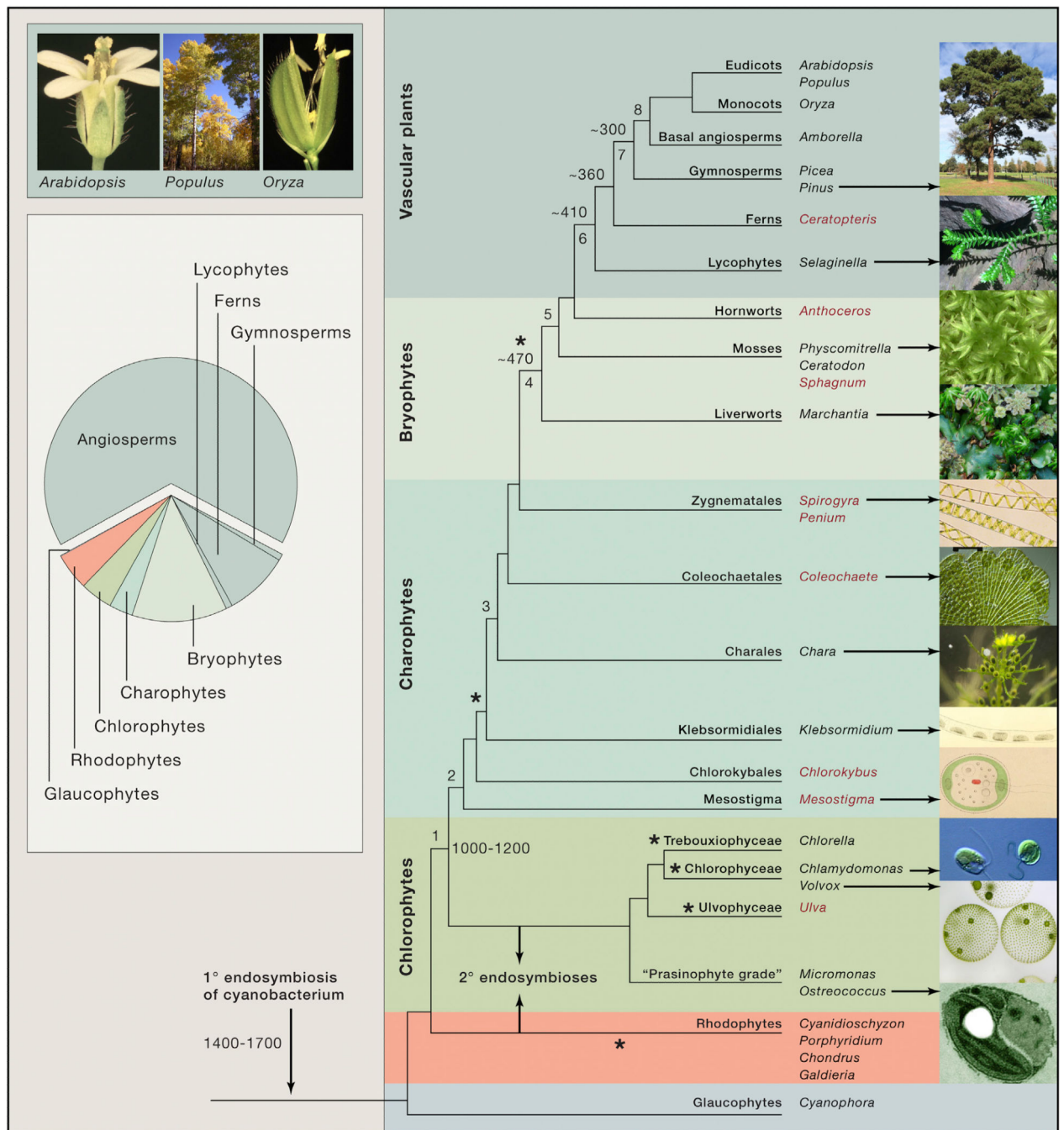
- Rushworth CA, Song B-H, Lee C-R, Mitchell-Olds T. *Boechera*, a model system for ecological genomics. *Mol. Ecol.* 2011; 20:4843–4857. [PubMed: 22059452]
- Sakakibara K, Ando S, Yip HK, Tamada Y, Hiwatashi Y, Murata T, Deguchi H, Hasebe M, Bowman JL. KNOX2 genes regulate the haploid-to-diploid morphological transition in land plants. *Science.* 2013; 339:1067–1070. [PubMed: 23449590]
- Sato S, Tabata S. *Lotus japonicus* as a platform for legume research. *Curr. Opin. Plant Biol.* 2006; 9:128–132. [PubMed: 16480917]
- Sato S, Nakamura Y, Kaneko T, Asamizu E, Kato T, Nakao M, Sasamoto S, Watanabe A, Ono A, Kawashima K, et al. Genome structure of the legume, *Lotus japonicus*. *DNA Res.* 2008; 15:227–239. [PubMed: 18511435]
- Schaefer DG, Zryd JP. Efficient gene targeting in the moss *Physcomitrella patens*. *Plant J.* 1997; 11:1195–1206. [PubMed: 9225463]
- Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, Pasternak S, Liang C, Zhang J, Fulton L, Graves TA, et al. The B73 maize genome: complexity, diversity, and dynamics. *Science.* 2009; 326:1112–1115. [PubMed: 19965430]
- Schwarz-Sommer Z, Davies B, Hudson A. An everlasting pioneer: the story of *Antirrhinum* research. *Nat. Rev. Genet.* 2003; 4:657–666. [PubMed: 12897777]
- Sessa EB, Banks JA, Barker MS, Der JP, Duffy AM, Graham SW, Hasebe M, Langdale J, Li F-W, Marchant DB, et al. Between two fern genomes. *Gigascience.* 2014; 3:15. [PubMed: 25324969]
- Sharma MK, Sharma R, Cao P, Jenkins J, Bartley LE, Qualls M, Grimwood J, Schmutz J, Rokhsar D, Ronald PC. A genome-wide survey of switchgrass genome structure and organization. *PLoS ONE.* 2012; 7:e33892. [PubMed: 22511929]
- Sharma S, Shrivastava N. Renaissance in phytomedicines-promising implications of NGS technologies. *Planta.* 2016; 244:19–38. [PubMed: 27002972]
- Shaw AJ, Schmutz J, Devos N, Shu S, Carrell AA, Weston DJ. Chapter Five – The *Sphagnum* Genome Project: A New Model for Ecological and Evolutionary Genomics. *Adv. Bot. Res.* 2016; 78:167–187.
- Shulaev V, Sargent DJ, Crowhurst RN, Mockler TC, Folkerts O, Delcher AL, Jaiswal P, Mockaitis K, Liston A, Mane SP, et al. The genome of woodland strawberry (*Fragaria vesca*). *Nat. Genet.* 2011; 43:109–116. [PubMed: 21186353]
- Singh R, Ong-Abdullah M, Low E-TL, Manaf MAA, Rosli R, Nookiah R, Ooi LC-L, Ooi S-E, Chan K-L, Halim MA, et al. Oil palm genome sequence reveals divergence of interfertile species in Old and New worlds. *Nature.* 2013; 500:335–339. [PubMed: 23883927]
- Somerville C, Koornneef M. A fortunate choice: the history of *Arabidopsis* as a model plant. *Nat. Rev. Genet.* 2002; 3:883–889. [PubMed: 12415318]
- Song J, Lu S, Chen ZZ, Lourenco R, Chiang VL. Genetic transformation of *Populus trichocarpa* genotype Nisqually-1: a functional genomic tool for woody plants. *Plant Cell Physiol.* 2006; 47:1582–1589. [PubMed: 17018558]
- Sørensen I, Fei Z, Andreas A, Willats WGT, Domozych DS, Rose JKC. Stable transformation and reverse genetic analysis of *Penium margaritaceum*: a platform for studies of charophyte green algae, the immediate ancestors of land plants. *Plant J.* 2014; 77:339–351. [PubMed: 24308430]
- Staton SE, Bakken BH, Blackman BK, Chapman MA, Kane NC, Tang S, Ungerer MC, Knapp SJ, Rieseberg LH, Burke JM. The sunflower (*Helianthus annuus* L.) genome reflects a recent history of biased accumulation of transposable elements. *Plant J.* 2012; 72:142–153. [PubMed: 22691070]
- Strable, J.; Scanlon, MJ. Maize (*Zea mays*): a model organism for basic and applied research in plant biology.; Cold Spring Harb. Protoc. 2009. p. emo1322009<http://dx.doi.org/10.1101/pdb.emo132>.
- Strain E, Hass B, Banks JA. Characterization of mutations that feminize gametophytes of the fern *Ceratopteris*. *Genetics.* 2001; 159:1271–1281. [PubMed: 11729168]
- Swaminathan K, Alabady MS, Varala K, De Paoli E, Ho I, Rokhsar DS, Arumuganathan AK, Ming R, Green PJ, Meyers BC, et al. Genomic and small RNA sequencing of *Miscanthus × giganteus* shows the utility of sorghum as a reference genome sequence for Andropogoneae grasses. *Genome Biol.* 2010; 11:R12. [PubMed: 20128909]



- Swaminathan K, Chae WB, Mitros T, Varala K, Xie L, Barling A, Glowacka K, Hall M, Jezowski S, Ming R, et al. A framework genetic map for *Miscanthus sinensis* from RNAseq-based markers shows recent tetraploidy. *BMC Genomics*. 2012; 13:142. [PubMed: 22524439]
- Szövényi P, Frangedakis E, Ricca M, Quandt D, Wicke S, Langdale JA. Establishment of *Anthoceros agrestis* as a model species for studying the biology of hornworts. *BMC Plant Biol*. 2015a; 15:98. [PubMed: 25886741]
- Szövényi P, Perroud P-F, Symeonidi A, Stevenson S, Quatrano RS, Rensing SA, Cuming AC, McDaniel SF. De novo assembly and comparative analysis of the *Ceratodon purpureus* transcriptome. *Mol. Ecol. Resour*. 2015b; 15:203–215. [PubMed: 24862584]
- Tam T, Catarino B, Dolan L. Conserved regulatory mechanism controls rooting cell development in land plants. *Proc. Natl. Acad. Sci. USA*. 2015; 112:3959–3968.
- Tang H, Krishnakumar V, Bidwell S, Rosen B, Chan A, Zhou S, Gentzittel L, Childs KL, Yandell M, Gundlach H, et al. An improved genome release (version Mt4.0) for the model legume *Medicago truncatula*. *BMC Genomics*. 2014; 15:312. [PubMed: 24767513]
- Tedder A, Ansell SW, Lao X, Vogel JC, Mable BK. Sporophytic self-incompatibility genes and mating system variation in *Arabidopsis alpina*. *Ann. Bot. (Lond.)*. 2011; 108:699–713.
- Thodey K, Galanie S, Smolke CD. A microbial biomanufacturing platform for natural and semisynthetic opioids. *Nat. Chem. Biol*. 2014; 10:837–844. [PubMed: 25151135]
- Thornton LE, Keren N, Ohad I, Pakrasi HB. *Physcomitrella patens* and *Ceratodon purpureus*, mosses as model organisms in photosynthesis studies. *Photosynth. Res*. 2005; 83:87–96. [PubMed: 16143910]
- Tomato Genome Consortium. The tomato genome sequence provides insights into fleshy fruit evolution. *Nature*. 2012; 485:635–641. [PubMed: 22660326]
- Tuskan GA, Difazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, Putnam N, Ralph S, Rombauts S, Salamov A, et al. The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science*. 2006; 313:1596–1604. [PubMed: 16973872]
- Van Hoeck A, Horemans N, Monsieus P, Cao HX, Vandenhove H, Blust R. The first draft genome of the aquatic model plant *Lemna minor* opens the route for future stress physiology research and biotechnological applications. *Biotechnol. Biofuels*. 2015; 8:188. [PubMed: 26609323]
- VanBuren R, Bryant D, Edger PP, Tang H, Burgess D, Challabathula D, Spittle K, Hall R, Gu J, Lyons E, et al. Single-molecule sequencing of the desiccation-tolerant grass *Oropetium thomaeum*. *Nature*. 2015; 527:508–511. [PubMed: 26560029]
- Vandenbussche M, Chambrier P, Rodrigues Bento S, Morel P. *Petunia*, your next supermodel? *Front. Plant Sci*. 2016; 7:72. [PubMed: 26870078]
- Vanneste K, Sterck L, Myburg AA, Van de Peer Y, Mizrahi E. Horsetails are ancient polyploids: Evidence from *Equisetum giganteum*. *Plant Cell*. 2015; 27:1567–1578. [PubMed: 26002871]
- International Peach Genome Initiative. Verde I, Abbott AG, Scalabrin S, Jung S, Shu S, Marroni F, Zhebentyayeva T, Dettori MT, Grimwood J, Cattonaro F, et al. The high-quality draft genome of peach (*Prunus persica*) identifies unique patterns of genetic diversity, domestication and genome evolution. *Nat. Genet*. 2013; 45:487–494. [PubMed: 23525075]
- Viaene T, Delwiche CF, Rensing SA, Friml J. Origin and evolution of PIN auxin transporters in the green lineage. *Trends Plant Sci*. 2013; 18:5–10. [PubMed: 22981345]
- Videvall E, Sletvold N, Hagenblad J, Ågren J, Hansson B. Strong maternal effects on gene expression in *Arabidopsis lyrata* hybrids. *Mol. Biol. Evol*. 2016; 33:984–994. [PubMed: 26685177]
- Wang R, Farrona S, Vincent C, Joecker A, Schoof H, Turck F, Alonso-Blanco C, Coupland G, Albani MC. PEP1 regulates perennial flowering in *Arabidopsis alpina*. *Nature*. 2009; 459:423–427. [PubMed: 19369938]
- Wang W, Haberer G, Gundlach H, Gläßer C, Nussbaumer T, Luo MC, Lomsadze A, Borodovsky M, Kerstetter RA, Shanklin J, et al. The *Spirodela polyrrhiza* genome reveals insights into its neoteneous reduction fast growth and aquatic lifestyle. *Nat. Commun*. 2014; 5:3311. [PubMed: 24548928]
- Wellman CH. The invasion of the land by plants: when and where? *New Phytol*. 2010; 188:306–309. [PubMed: 20941845]



- Wellman CH, Osterloff PL, Mohiuddin U. Fragments of the earliest land plants. *Nature*. 2003; 425:282–285. [PubMed: 13679913]
- Westfall PJ, Pitera DJ, Lenihan JR, Eng D, Woolard FX, Regentin R, Horning T, Tsuruta H, Melis DJ, Owens A, et al. Production of amorphadiene in yeast, and its conversion to dihydroartemisinin acid, precursor to the antimalarial agent artemisinin. *Proc. Natl. Acad. Sci. USA*. 2012; 109:E111–E118. [PubMed: 22247290]
- Wickett NJ, Mirarab S, Nguyen N, Warnow T, Carpenter E, Matasci N, Ayyampalayam S, Barker MS, Burleigh JG, Gitzendanner MA, et al. Phylotranscriptomic analysis of the origin and early diversification of land plants. *Proc. Natl. Acad. Sci. USA*. 2014; 111:E4859–E4868. [PubMed: 25355905]
- Wolf PG, Sessa EB, Marchant DB, Li FW, Rothfels CJ, Sigel EM, Gitzendanner MA, Visger CJ, Banks JA, Soltis DE, et al. An exploration into fern genome space. *Genome Biol. Evol.* 2015; 7:2533–2544. [PubMed: 26311176]
- Wos G, Willi Y. Temperature-stress resistance and tolerance along a latitudinal cline in North American *Arabidopsis lyrata*. *PLoS One*. 2015; 10:e0131808. [PubMed: 26110428]
- Wu CA, Lowry DB, Cooley AM, Wright KM, Lee YW, Willis JH. *Mimulus* is an emerging model system for the integration of ecological and genomic studies. *Heredity (Edinb)*. 2008; 100:220–230. [PubMed: 17551519]
- Wullschlegel SD, Weston DJ, DiFazio SP, Tuskan GA. Revisiting the sequencing of the first tree genome: *Populus trichocarpa*. *Tree Physiol*. 2013; 33:357–364. [PubMed: 23100257]
- Yang R, Jarvis DE, Chen H, Beilstein MA, Grimwood J, Jenkins J, Shu S, Prochnik S, Xin M, Ma C, et al. The reference genome of the halophytic plant *Eutrema salsugineum*. *Front. Plant Sci*. 2013; 4:46. [PubMed: 23518688]
- Yant L, Hollister JD, Wright KM, Arnold BJ, Higgins JD, Franklin FCH, Bomblies K. Meiotic adaptation to genome duplication in *Arabidopsis arenosa*. *Curr. Biol*. 2013; 23:2151–2156. [PubMed: 24139735]
- Yoon HS, Hackett JD, Ciniglia C, Pinto G, Bhattacharya D. A molecular timeline for the origin of photosynthetic eukaryotes. *Mol. Biol. Evol.* 2004; 21:809–818. [PubMed: 14963099]
- Young ND, Debellé F, Oldroyd GED, Geurts R, Cannon SB, Udvardi MK, Benedito VA, Mayer KFX, Gouzy J, Schoof H, et al. The *Medicago* genome provides insight into the evolution of rhizobial symbioses. *Nature*. 2011; 480:520–524. [PubMed: 22089132]
- Zang L, Zheng T, Chu Y, Ding C, Zhang W, Huang Q, Su X. Genome-wide analysis of the Fasciclin-Like Arabinogalactan protein gene family reveals differential expression patterns, localization, and salt stress response in *Populus*. *Front. Plant Sci*. 2015; 6:1140. [PubMed: 26779187]
- Zimin A, Stevens KA, Crepeau MW, Holtz-Morris A, Koriabine M, Marçais G, Puiu D, Roberts M, Wegrzyn JL, de Jong PJ, et al. Sequencing and assembly of the 22-gb loblolly pine genome. *Genetics*. 2014; 196:875–890. [PubMed: 24653210]



**Figure 1. Plant Diversity**

Depicted are relationships among the major lineages of plants: glaucophytes (freshwater algae), rhodophytes (red algae), and the Viridiplantae (chlorophytes, charophytes, and land plants [vascular plants and bryophytes]). Estimated dates for some nodes are listed in millions of years before present. The primary endosymbiotic event is estimated to have occurred at least 1.6 billion years ago. A deep split occurred within the green lineage creating the chlorophytes and the charophytes plus land plants. Note that both the charophytes and the bryophytes are grades and are not monophyletic. Key evolutionary

innovations are indicated at some nodes: (1) chlorophyll a and b, highly developed chloroplast grana; (2) lateral flagella; (3) phragmoplast, plasmodesmata, apical cell growth; (4) alternation of generations (embryo), sporopollenin, three-dimensional organization, archegonia and antheridia; (5) stomata; (6) vasculature (xylem and phloem), sporophyte dominant; (7) seeds, loss of archegonia and antheridia, pollen; and (8) flowers, carpels, outer integument. Evolution of simple multicellularity (differentiation largely limited to soma and gametes) and complex multicellularity (3-dimensional tissue differentiation) are indicated by \*. Taxa in red represent those whose genome sequencing are either in progress or proposed or reside at critical evolutionary positions (*Mesostigma*, *Coleochaete*). Secondary endosymbiotic events have occurred from both the red and green algal lineages. The relative species richness of the major clades is depicted in the pie chart. The vast majority of species within the Plantae are angiosperms (350,000 species), with other groups having substantially fewer described species (numbers approximated): glaucophytes, 21; rhodophytes, 7095; chlorophytes, 6263; charophytes, 4644; bryophytes, 18,200 (liverworts, 7,270; mosses, 11,000; hornworts, 215); lycophytes, 1225; ferns, 12,000; and gymnosperms, 1000. Photos from top: *Pinus radiata*, *Selaginella kraussiana*, *Physcomitrella patens*, *Marchantia berteroana*, *Spirogyra* sp., *Coleochaete orbicularis*, *Chara* sp., *Klebsormidium flaccidum*, *Mesostigma viride*, *Chlamydomonas reinhardtii*, *Volvox* sp., and *Ostreococcus tauri*. Thanks to individuals who generously provided photos of poplar (Gayle Dupper, Institute of Forest Genetics, Placerville), *Physcomitrella* (Keiko Sakakibara), *Spirogyra* (Bram Van de Poel, University of Maryland), *Coleochaete*, *Chara* (Charles Delwiche, University of Maryland), *Chlamydomonas* (James Umen, Salk Institute), *Volvox* (Frank Fox, [www.mikrofoto.de](http://www.mikrofoto.de)), and *Ostreococcus* (Hervé Moreau, Université Pierre et Marie Curie-Paris). Other photos by J. L. B. *Mesostigma* image from (Lauterborn, 1898) and *Klebsormidium* image from (Klebs, 1896). Figure adapted from Bowman et al. (2007).

Table 1

Some Existing and Emerging Plant Models

| Classification     | Family       | Species Name  | Genome Size (Mb) <sup>1</sup> | Genome Sequence Status | Properties   | References  |
|--------------------|--------------|---|-------------------------------|------------------------|--|---|
| Angiosperm eudicot | Brassicaceae | <i>Arabidopsis thaliana</i> (mouse-ear cress)   | 135                           | complete               | predominant plant model system                     | Provart et al., 2016  |
|                    |              | <i>Arabidopsis halleri</i>  | 145.5                         | nearly complete        | heavy metal tolerance/accumulation                 | Kubota et al., 2015; Kosugi et al., 2015                      |
|                    |              | <i>Arabidopsis lyrata</i> (lyrate rock cress)   | 207                           | nearly complete        | self-incompatible <i>A. thaliana</i> relative      | Hu et al., 2011; Joschinski et al., 2015; Wos and Willi, 2015 |
|                    |              | <i>Eutrema salsugineum</i> formerly known as <i>Thellungiella halophila</i> (saltwater cress) | 238.5                         | nearly complete        | salt-tolerant <i>A. thaliana</i> relative          | Inan et al., 2004; Yang et al., 2013                          |
|                    |              | <i>Cardamine hirsute</i> (bitter cress)   | 549 <sup>2</sup>              | in progress            | compound leaves                                    | Hay et al., 2014; Bar and Ori, 2015                           |
|                    |              | <i>Lotus japonicas</i> (birdsfoot trefoil)  | 472 <sup>3</sup>              | in progress            | nitrogen fixation                                  | Sato and Tabata 2006; Sato et al., 2008                       |
| Fabaceae           |              | <i>Medicago truncatula</i> (barrel medic)   | ~454–526 <sup>4</sup>         | nearly complete        | nitrogen fixation                                  | Young et al., 2011; Tang et al., 2014                         |
|                    |              | <i>Eucalyptus globulus</i> (blue gum)   | 530 <sup>5</sup>              | nearly complete        | wood, lignification, biofuel                       | Araújo et al., 2014; Myburg et al., 2014                      |
|                    |              | <i>Eucalyptus grandis</i> (rose gum)  | 640 <sup>5</sup>              | nearly complete        | wood, biofuel                                      | Myburg et al., 2014   |
| Orchidaceae        |              | <i>Phalaenopsis</i> spp. (moth orchid)  | ~1160 <sup>6</sup>            | nearly complete        | complex flower development, commercial flower crop | Hsu et al., 2011; Pan et al., 2014; Cai et al., 2015          |
| Phrymaceae         |              | <i>Mimulus guttatus</i> (sheep monkeyflower)  | 322                           | nearly complete        | ecological studies, flower evolution               | Wu et al., 2008   |
| Ranunculaceae      |              | <i>Aquilegia caerulea</i> (Colorado blue columbine)   | 306                           | nearly complete        | flower evolution and ecology                       | Kramer and Hodges, 2010                                       |

| Classification     | Family     | Species Name  | Genome Size (Mb) <sup>1</sup> | Genome Sequence Status | Properties                                     | References   |
|--------------------|------------|---|-------------------------------|------------------------|--|--|
|                    | Rosaceae   | <i>Fragaria vesca</i> (wild strawberry)             | 240                           | complete               | fruit development                              | Shulaev et al., 2011                               |
|                    |            | <i>Prunus persica</i> (peach)                       | 225                           | nearly complete        | fruit tree development                         | International Peach Genome Initiative et al., 2013 |
|                    | Salicaceae | <i>Populus trichocarpa</i> (black cottonwood)       | ~422.9                        | complete               | wood   | Tuskan et al., 2006; Wulfschleger et al., 2013     |
|                    | Solanaceae | <i>Solanum lycopersicum</i> (tomato)                | ~900                          | complete               | fruit ripening                                 | Tomato Genome Consortium, 2012                     |
|                    |            | <i>Petunia × hybrid</i> (petunia)                   | ~1,300                        | in progress            | flower development, evolution                  | Vandenbussche et al., 2016                         |
|                    | Vitaceae   | <i>Vitis vinifera</i> (grape)                       | 487                           | complete               | wine   | Jaillon et al., 2007; Pulvirenti et al., 2015      |
| Angiosperm monocot | Araceae    | <i>Lemna minor</i> (common duckweed)                | 472 <sup>7</sup>              | nearly complete        | aquatic plant for phytoremediation and biofuel | Cantó-Pastor et al., 2015                          |
|                    |            | <i>Spirodela polyrrhiza</i> (greater duckweed)      | 158 <sup>8</sup>              | in progress            | aquatic plant for phytoremediation and biofuel | Wang et al., 2014                                  |
| Poaceae            |            | <i>Brachypodium distachyon</i> (purple false brome) | ~272                          | nearly complete        | non-crop grass model                           | Brkljačić et al., 2011                             |
|                    |            | <i>Miscanthus sinensis</i> (Chinese silver grass)   | 5,500 <sup>9</sup>            | in progress            | bioenergy plant                                | Ma et al., 2012; Swaminathan et al., 2012          |
|                    |            | <i>Oropetium thomaeum</i>                           | 245 <sup>10</sup>             | nearly complete        | drought model                                  | VanBuren et al., 2015                              |
|                    |            | <i>Oryza sativa</i> (rice)                          | 372                           | complete               | major crop, grass model                        | International Rice Genome Sequencing Project 2005  |
|                    |            | <i>Panicum virgatum</i> (switchgrass)               | ~1,500 <sup>11</sup>          | nearly complete        | bioenergy crop                                 | Sharma et al., 2012                                |
|                    |            | <i>Setaria viridis</i> (green foxtail)              | ~395                          | nearly complete        | non-crop grass                                 | Bennetzen et al., 2012                             |
|                    |            | <i>Sorghum bicolor</i> (sorghum)                    | ~732                          | complete               | potential bioenergy crop                       | Paterson et al., 2009                              |

| Classification   | Family             | Species Name  | Genome Size (Mb) <sup>1</sup> | Genome Sequence Status | Properties                            | References                                   |
|------------------|--------------------|---|-------------------------------|------------------------|---------------------------------------|--|
| Basal angiosperm | Amborellaceae      | <i>Amborella trichopoda</i>                         | 748                           | nearly complete        | flowering plant evolution             | Amborella Genome Project (2013)              |
| Gymnosperm       | Pinaceae           | <i>Picea abies/ Picea glauca</i> (spruce)           | ~20,000 <sup>13</sup>         | in progress            | wood                                  | Nystedt et al., 2013; Biröl et al., 2013     |
|                  |                    | <i>Pinus taeda</i> (loblolly pine)                  | ~22,000 <sup>14</sup>         | nearly complete        | wood                                  | Zimin et al., 2014                           |
| Pteridophyta     | Pteridaceae        | <i>Ceratopteris trichardii</i> (triangle waterfern) | 11,260 <sup>15</sup>          | in progress            | model fern                            | Strain et al., 2001; Atallah and Banks, 2015 |
|                  | Azollaceae         | <i>Azolla filiculoides</i> (water fern)             | 740 <sup>16</sup>             | nearly complete        | model fern, nitrogen fixation         | Qiu and Yu, 2003; Sessa et al., 2014         |
| Lycopodiophyta   | Selaginellaceae    | <i>Selaginella moellendorffii</i> (spikemoss)       | ~106 <sup>17</sup>            | complete               | clubmoss evolution                    | Banks et al., 2011                           |
| Anthocerotophyta | Anthocerotaceae    | <i>Antheroceros agrestis</i> (field hornwort)       | 83 <sup>18</sup>              | in progress            | hornwort evolution                    | Szövényi et al., 2015a                       |
| Bryophyta        | Ditrichaceae       | <i>Ceratodon purpureus</i> (fire moss)              | 340 <sup>19</sup>             | in progress            | photosynthesis, heavy metal tolerance | Thomton et al., 2005                         |
|                  | Funariaceae        | <i>Physcomitrella patens</i>                        | 473                           | complete               | moss evolution, moss development      | Knight et al., 2009                          |
| Marchantiophyta  | Marchantiaceae     | <i>Marchantia polymorpha</i>                        | 226                           | in progress            | liverwort evolution                   | Ishizaki et al., 2016                        |
| Charophyta       | Klebsormidiaceae   | <i>Klebsormidium flaccidum</i>                      | ~104 <sup>20</sup>            | nearly complete        | filamentous freshwater alga           | Hori et al., 2014                            |
| Chlorophyta      | Chlamydomonadaceae | <i>Chlamydomonas reinhardtii</i>                    | ~120 <sup>21</sup>            | complete               | motile unicellular freshwater alga    | Harris, 2001; Merchant et al., 2007          |
|                  | Mamiellaceae       | <i>Ostreococcus tauri</i>                           | 12.56 <sup>22</sup>           | complete               | miniscule unicellular marine alga     | Courties et al., 1994; Derelle et al., 2006  |
|                  | Volvocaceae        | <i>Volvox carteri</i>                               | 131.2                         | nearly complete        | multicellular freshwater alga         | Prochnik et al., 2010                        |

<sup>1</sup> <https://phytozome.jgi.doe.gov/pz/portal.html> unless otherwise noted.



<sup>2</sup> [http://www.brassica.info/resource/sequencing/bmap/CSP2012Proposal\\_Version11\\_Final.pdf](http://www.brassica.info/resource/sequencing/bmap/CSP2012Proposal_Version11_Final.pdf).

<sup>3</sup> Sato et al., 2008.

<sup>4</sup> Young et al., 2011.

<sup>5</sup> Myburg et al., 2014.

<sup>6</sup> Cai et al., 2015.

<sup>7</sup> Van Hoeck et al., 2015.

<sup>8</sup> Wang et al., 2014.

<sup>9</sup> Rayburn et al., 2009; Cichorz et al., 2015.

<sup>10</sup> VanBuren et al., 2015.

<sup>11</sup> Sharma et al., 2012.

<sup>12</sup> Schnable et al., 2009.

<sup>13</sup> Nystedt et al., 2013; Birol et al., 2013.

<sup>14</sup> Zimin et al., 2014.

<sup>15</sup> Sessa et al., 2014.

<sup>16</sup> <http://www.azollagenome.net/>.

<sup>17</sup> Banks et al., 2011.

<sup>18</sup> Szövényi et al., 2015a.

<sup>19</sup> Thornton et al., 2005.

<sup>20</sup> Hori et al., 2014.

<sup>21</sup> Merchant et al., 2007.

<sup>22</sup> Derelle et al., 2006.